



#5

OIEP

ENTERED

RAW SEQUENCE LISTING

DATE: 03/14/2002

PATENT APPLICATION: US/09/877,935

TIME: 14:17:34

Input Set : A:\13294-002001.txt

Output Set: N:\CRF3\03142002\I877935.raw

```

3 <110> APPLICANT: Pinto, Daniel
4     Robine, Sylvie
5     Jaisser, Frederic
6     Louvard, Daniel
8 <120> TITLE OF INVENTION: REGULATORY SEQUENCES OF THE MOUSE VILLIN GENE - USE IN
TRANSGENESIS
10 <130> FILE REFERENCE: 13294-002001
12 <140> CURRENT APPLICATION NUMBER: US 09/877,935
13 <141> CURRENT FILING DATE: 2001-06-08
15 <150> PRIOR APPLICATION NUMBER: PCT/EP 98/08009
16 <151> PRIOR FILING DATE: 1998-12-09
18 <160> NUMBER OF SEQ ID NOS: 20
20 <170> SOFTWARE: PatentIn Ver. 2.1
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 8995
24 <212> TYPE: DNA
25 <213> ORGANISM: Mus musculus
27 <220> FEATURE:
28 <221> NAME/KEY: intron
29 <222> LOCATION: (3489)..(8981)
31 <220> FEATURE:
32 <221> NAME/KEY: exon
33 <222> LOCATION: (3443)..(3487)
34 <223> OTHER INFORMATION: exon 1
36 <400> SEQUENCE: 1
37 gatctggtgc accaaggaca ctgtggtccc agcactgggg aggtggaggg aggaggggtca 60
38 gaagttaaag gtcacacctg gttacatagc aaggtttcag ccagcttcag ctacatgaaa 120
39 cctttgtttg ttgtttgttt tgttttaaag cattaataaa taataccata aggaggttgg 180
40 cagtgggtggc agacaccttt aattccagta ttcaggaggc agaagcaggc agatctctgt 240
41 gagttcgaag tcagcctagt ctgcaaagct agttccagga tggcaagggc tacacagaga 300
42 aaccttgtct cataaaacca aagtagtagt agtagtagta atgccataga gaaaatttga 360
43 gtccattcag gatggaccat cctataagat gattctcttg acccaggtaa gctaattgtca 420
44 tggggaaagg ggatgggact gtcctagatt aaaaagtgtc gaggcgatgc ctattctcaa 480
45 ttgtattcca tatgaaaagg ctgataaggc ccaagagaag tggaaactggg actctggact 540
46 gaagacgtga cggccttata aacactggca cttataaaca cttataaaca ctggcacagg 600
47 cgttcagggt tgaagatcac ttccaacca cagaacagaa agtgctcgct cgtcctcagc 660
48 gtagcgagca ctggctgcag aagagtgata tttagtgaat gctaccttca caatatcttt 720
49 gcacttatca catacacgtg tcaaatgtgc taactcccta gtccacagat ggctgttaca 780
50 ctggtttctg ctttcccatc tggttgacat ttgtcagaac cagaaattag aaatgtgggt 840
51 atttatttgt gtgtgagga caccatccag ggcttttcac atttcaggca catggtttac 900
52 taactgggct acttctccaa cggtttgaaa ccatttgttt tatatttact tatttttgtt 960
53 gcatgaggta ggcattgata cgtatgtata ggagtcagtc atgtggctgc taccctcaaa 1020
54 atcattgcag atccccagca agtgaagtca ccgagcgttg taagttgtta tgtgggactg 1080
55 ggagccaagg ctgggttctc tgcaagagca gccagtggcc ttaacatgg gaccagctct 1140

```

RAW SEQUENCE LISTING

DATE: 03/14/2002

PATENT APPLICATION: US/09/877,935

TIME: 14:17:34

Input Set : A:\13294-002001.txt

Output Set: N:\CRF3\03142002\I877935.raw

```
56 ctaggcctaa ggtaatcttt agtttttttaa aaatatatat tctcagccgg gtgtggtggc 1200
57 acacgccttt aatcccagca cttagagaggc tgaggtgtag gaattataca cacaggccag 1260
58 ctgggggtgca gagcttggcc ctgtttttttt tgttttttct ttatgtgcac tgggtgtctta 1320
59 cctgcgtgta tgtccgtgca aggggtgtcag atcccttggg gctggagtta aagacagttg 1380
60 tgatcacgct gccgttacag atgctggaaa ttgaaccagc gtgtccctag agaagcagcc 1440
61 agtgcctctta acttctgagc cacccttcca accctgcttt tagagactct taaccttttg 1500
62 tgtaatgtgg gaactgagtg gatcttgcac ttaccaagtg tgtgctgcgc tgtagcatca 1560
63 ctgagcccggt acccacacga ctagtggata cagttaaagg gcaaacactt aacaatgaca 1620
64 atagtgggat agagtttgaa tatagtcctg agctattggg tagcgtgacc tttgctgtcc 1680
65 ttagcatgtg ctgtgagaag atagaaaaat gaagacttga gtctagtccct ggaaccaca 1740
66 gaggcaggcg agaaccact cctgaaagtt gttctctgag ctccacatac aacttcacat 1800
67 aatagttaca atgataataa taattagtaa attcttttaa aaggatatatg ttgggagggg 1860
68 gagatggctc agcttccagg agcacttgct gctcttgagc aggacctaga ttcagttccc 1920
69 aggactcata tgggtggctca cagccatctg taaatccagt tccagagggg tccacaccct 1980
70 cttctggcct ccacaggcac cacatacata gtacacagac atacatgcag gcaaaacacc 2040
71 catacacaca taaataaata aggaaactta aaaggtgcat gtgttggtta acattgtgct 2100
72 tacacatgct gattgaagac atgtacaacg cacacactga agagggatct ggggctggag 2160
73 agatggctca gcggttaaga gcactgactg ctcttccgaa ggaaggctct gagttcaaat 2220
74 cctagcaacc acatggtggc tcacaacat ccatatgag atctgacacc ctcttctggg 2280
75 gcatctgaag acagctgcag agctacagtg tacttagata tactaataaa taaatctttt 2340
76 tttaaaaaaa tgaagaggga tctgagacac ctcaaaagag attatgagca gtgactcacg 2400
77 ggtgattatc tatcctggag tttttccttt ccgcttggct tgcaactggg tggacagacg 2460
78 ccccttttca ttcacaagaa cgggtgctac attatttctg aacaaaacag cacctgcagt 2520
79 atgtttactg tccttgctga ctatgagcac gcgcacgcgc gcgcgcacac acacacacac 2580
80 acacacacac acacacacac acacacacac attcagtcct cagagctctt ggggaaggtca 2640
81 agaagaggct gccctcaaac acgatcttca tctttccctc cttaaaggaga ccacgattcc 2700
82 aagggtggcag aagatctaca gggggcagag gcagggaggg ggaagcaggc catggtttcc 2760
83 agagacctac agcagagggc agcaaggcag atccccaggt ccagggcagg gaggtggagg 2820
84 cccttggtcc gagggagaag caggcggcag aacaggggtc aaaggcacag gtttatggca 2880
85 gctcataaaa gtggagggtcg tggctcactc agaaaggagg agaaggggaa aggcccttgt 2940
86 gccactgag cgaggggtcat gctgagtagg agagatctgc aggggtgcca ggagccccac 3000
87 ctgtctgtcc caagggaacc ccaagtgtga actctggcct tgggtgctga gttccagcta 3060
88 caagacccca ggagtcctac tccatcccca tccagtgcct cctcgccccg ccacacccca 3120
89 ccccgactc ccgtgccact tctctagggc tggaggggtg ccagccctgg tgggggttgc 3180
90 ctacctgcag gtagagccca ggtcctagcc ggaagtgcac cccatccctg aagctgcaga 3240
91 gccaagggcg gggcacacgg cagctcagcc tgtcaggctg ttgctgggct ctaggttccc 3300
92 agggacctgg gaacctactt cccacacccc ccatccatc tctctggggc cctatcttcc 3360
93 cttatatggt gaaggaaagt cctggggggg ggggggtggt gtgaggacaa aggtcgttgc 3420
94 gtctcctgca gccagcttgc ca caa ctt cct aag atc tcc cag gtg gtg gct 3472
95 gcc tct tcc aga cag gtaaggcaat tgggtgggga cacatggtga ccacaggtg 3527
96 ttggagggga cagggtcctt gcttctctct ggcagcctgt gctttctgta gcaccttgg 3587
97 ataagtttgg gggtagggtg aggtgctctg aaactctgaa agaagcaaga agccagcagg 3647
98 ctgtcttggg ccttcaatga aggaagttca cagacccctt tctctgtaag tcaccttgc 3707
99 tcatctgtg tagattccct gggaccaagg tggctcctgg gactcagatt tctacaatta 3767
100 aaatcaggac agtccctgaga cttggactcc gtgcctgtat ttactacttc tctctggctg 3827
101 ctcatctctg tgttcatgtc ttacacatct gaaatggttt ctttgtgtca ccattcccct 3887
102 gacactcctg ggaggtcgta tccttggcac atgtatcctg ggatgtaagc tgcagccacc 3947
103 aggagagagg gggagagtca ggagctgtgt cctaggccct attaggcctg gacatcacc 4007
104 ctttccctaga aatggccctt ccatttttct gttacatga tctattttat atcagagtgg 4067
```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/877,935

DATE: 03/14/2002

TIME: 14:17:34

Input Set : A:\13294-002001.txt

Output Set: N:\CRF3\03142002\I877935.raw

```

105 gcagtgaagg ccaaaccctgc ccagaagttt gggactcact cagaccaagg ttatctgctc 4127
106 agaaatcccc ctgtcacttg aggttgggag aatctgcctc tgggggcttc caggtcttgg 4187
107 ttagcaggag ggtatccttt gtatagggca tgacctagtc tatggtgtta ctacattcct 4247
108 gtccagttaa aagctggaac taaaaccac ggcagcgccc aggattctct acagttgtac 4307
109 cccaagaaca acaagacagt agatatgcaa ggataggtag ctggggagaa gaagaactta 4367
110 aaccccccca aaggcccaca ggttccgttc cctagttcac aatgccagta tgagtgtag 4427
111 ctactatggg ctgtgagttg gtagctacaa gcatgagtga tgttcatgtg tgtagtgtgt 4487
112 ataactctgag cacttgggag gctgaagcag gaggattgct atatgtttga ggccagcctg 4547
113 agctatagag cgagactttg tctttaagaa aaaaatgaaa gccagcagc ggtggcacac 4607
114 gcctttaatc ccagcacttg ggaggcagaa gcaggcagat ttctgagttc aaggccagcc 4667
115 tggctctatag agtgagttcc aggacagcca gggctacaca gagaaaccct gttttgaaaa 4727
116 accagaaaaa caaaaacaaa caaaacaaa caaaacccaa acccaaacc aaacctctca 4787
117 tctctcatct ctctaggctg tgtctgtcta ggtggtagag tttggggact tcagacttat 4847
118 atataaatag gcccttttat cactggtcag agacgagaaa ggtttcagtc tgggacacag 4907
119 tgggaccctg agaaagtact ccttgccagc caaaaattc tgggaaggct tcctggagga 4967
120 agtgtgtccc gatcagacta ctgttctaga aggcagaaga gaggggttga agaattgttg 5027
121 tggacagaca gttggaacag aaggacagga gggggaggca tccaagattc tgaacatgta 5087
122 gctgactttt ggttctctgg gtgacaagtg tccccagggt atagggtgtg agaaagggga 5147
123 ccagggggtga gccaatgagt tcaagttgag ggacacatcc agcccagggt cctgtctggc 5207
124 aagctaaaga atgagagccc tctaaccctc cctgaagttt aggggagaca ggagagctga 5267
125 ggagatcctt ctagggtgaa ggagaggtat ctgctctgac caacatggct aggagcagaa 5327
126 gcagttggac cagttacccc tcagaaccag ccatcccctc ttggctctaa ggaggctggg 5387
127 cccctttctg ttaagaatc ttacttttct tcagagagag gcagcaagcc tttgtcccct 5447
128 ccctgttggt caataaacac ccctgtgtgt aacattagtt tattttactg tcagtttctg 5507
129 ccaggacagt ccatctggta gacctctgct cctaactcac caaggtatgg cccacattcc 5567
130 tcaccagaaa gagtgcagaa gagagcctta gagaaagggt aacagtaaca aagatggcca 5627
131 gaataaaaca aaaactacta tcctttgtac ccaaatgggt tttgtgaac caggaggggg 5687
132 tgtgtgagtg tatgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 5747
133 tgtgtgtgtg tgtcttgggg gacttttcat gctaaagaat atctgatatt ggcgcccatg 5807
134 ccaacagggg tattggggag agtcaggctt ctgcaaacac agtaagctgc ccaagatgga 5867
135 ttggtggcct gaatcaccaa ggggcaggct gatcagagtg gacagaacat cacaagataa 5927
136 gccaccctgt ggggtcaga agaggaggtt tacaagaggt aaaggccaag ccatttatta 5987
137 tccaagacat gactcaaaat caaagtcaa ggagagatta gctggagaga tggggctgtc 6047
138 agtgtgggac acctgacctt gcacttatta gtcactaggc caaggagcag tcacagaggg 6107
139 tgactgggtc ctactcagct tggagcaggg acgtggagaa tgggtgacct ccactctgat 6167
140 ggagaggggt gagcaccacc aggtacaagt gttccctgtg tctcatgcca ggattcctgg 6227
141 ccagttttca aaggactaag gactcatctc tgggtgaaac aaagtatcca agccctaagc 6287
142 cccatttttg tctaattaaa tcagaacccc tggggatgca ggctctgagc agcaggagct 6347
143 ttttaaaaag ctcccagggt attctgatca gcagctggaa caaacacagc tacaggttca 6407
144 aacagaaaaga ggcaaagcta gggaaagctt gggatgggga gccttcttcc aggccagtag 6467
145 atggaggctg gtttagcagtg gtggcagctt ctctctgctt gtcatatagc tatccatcca 6527
146 ctcatccatc catacaccca cccatccatt tatgcaccca tcttccatc catccatcta 6587
147 tccagctacc caccacgca tccatccaaa ccttcccttt ccttctcttt ctttcttttt 6647
148 tcttctactc attcatttat ccaacagaga actggtattg tactaaatgt gggagattta 6707
149 attaatTTTT agaagctctg ttgattgact gattgtgcat gtatgtggac aggtacatac 6767
150 cacagcacac gtgtggcaat cggagaaagg ttttgggtgt tgttttctct tcccaccgtg 6827
151 tgggttctgg ggattgaact caaattatcg ggctgggtgg aagtgtcttt accaccgagc 6887
152 cattttctg acacatcatt attattagaa agcatcttat gtagtccagg ctggcctcaa 6947
153 gcttctatg tcgccacgga tgacctttaa ctctgtctt tccagcctcc acccgagtgc 7007

```

RAW SEQUENCE LISTING

DATE: 03/14/2002

PATENT APPLICATION: US/09/877,935

TIME: 14:17:34

Input Set : A:\13294-002001.txt

Output Set: N:\CRF3\03142002\I877935.raw

```

154 taggtttaca ggtgttcaac tggatgaatgc ctttaatccc agcactctgt ggggggggggg 7067
155 ggggaggcgg atccctgagt tggaggccag tttggtctac agagtttcag gatacctggg 7127
156 gctatacagg gaaaccctat cccaaacaaa caaacaacaa aacaaaaaat attctgtgca 7187
157 ataatacag agattagagg atattagtag ggtagtaggg ctggtgaggg agagtcatgc 7247
158 tttcttttgt attataatag taaagtactc acaagatgca ttatctatct atctatctat 7307
159 ctatctatct atctatctat ctatctacct acctacctac ctatccatcc atccatctat 7367
160 cgtatagccc aggctgcttt gactctgaat gctcctatct ctgggtcaac tcttcacccc 7427
161 tagtggtggg tttaccaaca cccagacatt tattttatct tgttttatct tattaatcta 7487
162 ggagctcagg gtgggactca gggctctgtg catgctaagc aagctctctg ccacagagct 7547
163 gcagctccag tccccatttt gttcagggtga ctctgtgaca gttgtcatat tgcagcgt 7607
164 atgtagctct ctccacctcc cagttccagc actttctggt catcccagtg ggcgggcaac 7667
165 tctgtgctca ccagtgccct gttccctgtc ttcagacctc catatttgcc tgtctgaaca 7727
166 gttcatgtaa atgggatgcg ttcctgtgta ttcttttatg gctggccccct ttatcttagc 7787
167 acagtttgtg ttgggccatg tgtcaactgt ataactatct ttatcatcat cttatggctt 7847
168 aatagtgttc ctttgtgtgg ataaaccact ttctgtttca tttactgatg gaaatttgtg 7907
169 gccccacccc cacccttttt ttttttatct gagacaaggc ctttctgtgt aatcttgcaa 7967
170 tcttggtgtg cctgagctca ctctgtagac caggctgtga ggctgtcctt ccacttttga 8027
171 cactcctgtg aacagagtag ccatgaactt caaagacaat tttctgtttt ggtttgtttt 8087
172 ttacatttgt gtgtgtatgc gtgtatatgt gcatgtttgt gtcttcaggc gctcacatgt 8147
173 gtgtacctgt gtgtgggaca gagaacaaac cgatgtgcca ttcttcagat actacgcctc 8207
174 ttgttaatat gtatgtatta tgtatgttta tttagtgtgc ccaagtatgc aggtattttg 8267
175 ttggagtttt caccttcctt tgtgggctct ccgcattaaa ctgagctcct cgggctagt 8327
176 agcaatgcct tcaactgatg agccatctcg ctgccccctg tgccacctcc tcttatcttc 8387
177 ccagatggga ctacgcactg cactggccta aagctcacca agtcatccag agtggctagc 8447
178 caggagact cagggatag ctggcctctg cctccacagt gctagaatta caggcataca 8507
179 tcaactgctg aagattttta acctgaatcc tgaggataga gcaggcactc taccaatgga 8567
180 gggttctttt tgtgttttgt ttggtttcct ctgcataaga tcaggcagtc tgaaatagt 8627
181 tagcctgggc tacataacat ctgtctcaa aaagcctata gaggtaggga ggtcgaggct 8687
182 aaagaagagc cttaagccgg ctgtgatagc acacaggata gcctgcacta tatagcaaga 8747
183 cctgttttca aaaacatgga gggaggggta tgttttaagt gctgggctgt gtaacaggca 8807
184 ctaaggagc caatgtagac atttgactaa gaaaggatca tcatcaaagc cgggtgggca 8867
185 gggtagaggt tggactacag tggatcaagc ccccatagga agccagtttc ctttcttct 8927
186 ctgggctca agcctggctc gacggccact gctctcacat gccttctcct ctaggctcgt 8987
187 ccaccatg 8995

```

189 <210> SEQ ID NO: 2

190 <211> LENGTH: 23

191 <212> TYPE: DNA

192 <213> ORGANISM: Artificial Sequence

194 <220> FEATURE:

195 <223> OTHER INFORMATION: oligonucleotide

197 <400> SEQUENCE: 2

198 gagtggatg gttgagagag cct

23

200 <210> SEQ ID NO: 3

201 <211> LENGTH: 18

202 <212> TYPE: DNA

203 <213> ORGANISM: Artificial Sequence

205 <220> FEATURE:

206 <223> OTHER INFORMATION: oligonucleotide

208 <400> SEQUENCE: 3

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/877,935

DATE: 03/14/2002

TIME: 14:17:34

Input Set : A:\13294-002001.txt

Output Set: N:\CRF3\03142002\I877935.raw

```

209 catagttctc gttccggt 18
211 <210> SEQ ID NO: 4
212 <211> LENGTH: 44
213 <212> TYPE: DNA
214 <213> ORGANISM: Artificial Sequence
216 <220> FEATURE:
217 <223> OTHER INFORMATION: oligonucleotide
219 <400> SEQUENCE: 4
220 gatctcccag gtggtggctg cctcttcag acaggctcgt ccac 44
222 <210> SEQ ID NO: 5
223 <211> LENGTH: 44
224 <212> TYPE: DNA
225 <213> ORGANISM: Artificial Sequence
227 <220> FEATURE:
228 <223> OTHER INFORMATION: oligonucleotide
230 <400> SEQUENCE: 5
231 catggtggac gagcctgtct ggaagaggca gccaccacct ggga 44
233 <210> SEQ ID NO: 6
234 <211> LENGTH: 18
235 <212> TYPE: DNA
236 <213> ORGANISM: Artificial Sequence
238 <220> FEATURE:
239 <223> OTHER INFORMATION: oligonucleotide
241 <400> SEQUENCE: 6
242 caacttccta agatctcc 18
244 <210> SEQ ID NO: 7
245 <211> LENGTH: 20
246 <212> TYPE: DNA
247 <213> ORGANISM: Artificial Sequence
249 <220> FEATURE:
250 <223> OTHER INFORMATION: oligonucleotide
252 <400> SEQUENCE: 7
253 attcaggctg cgcaactgtt 20
255 <210> SEQ ID NO: 8
256 <211> LENGTH: 24
257 <212> TYPE: DNA
258 <213> ORGANISM: Artificial Sequence
260 <220> FEATURE:
261 <223> OTHER INFORMATION: oligonucleotide
263 <400> SEQUENCE: 8
264 gcaacagtcg ctggacatca cagg 24
266 <210> SEQ ID NO: 9
267 <211> LENGTH: 20
268 <212> TYPE: DNA
269 <213> ORGANISM: Artificial Sequence
271 <220> FEATURE:
272 <223> OTHER INFORMATION: oligonucleotide
274 <400> SEQUENCE: 9
275 ccacggacaa ctgcgttgat 20

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/877,935

DATE: 03/14/2002

TIME: 14:17:35

Input Set : A:\13294-002001.txt

Output Set: N:\CRF3\03142002\I877935.raw